2.1 Foundation and Model Structure

- One example.

  One shoe company asks you to look for the average foot size of Chinese in China. It is clear that the sample space is all people in China. You can first get an i.i.d. sample and calculate its average, say 38cm. So you may claim that the population mean roughly equals to 38cm.

  If someone doesn't believe this, he/she can get another i.i.d. sample and calculate the average. This new average should be almost 38cm as long as the sample size is large enough and the sample design is almost the same.

  That is, 38cm is a convincing number as the average foot size of Chinese.

- Space

  We go back to the example of HSI in Section 1. Let r_t be the daily log-return of HSI at day t.

  r_t has its own probability space $(\Omega_t, \mathcal{F}_t, \mathcal{P}_t)$, and

  $$m_t \equiv E r_t = \int_{\Omega_t} r_t d\mathcal{P}_t.$$

  On day t, we have only one observation.

  How to estimate m_t?

  We use the historical data, say $r_1, \cdots, r_n$, to estimate m_t, i.e.,

  $$m_t \approx \frac{1}{n} \sum_{t=1}^{n} r_t.$$

  Do you believe it?

First, we need to assume that

$$(\Omega_t, \mathcal{F}_t, \mathcal{P}_t) \equiv (\Omega, \mathcal{F}, \mathcal{P})$$

—time invariant, i.e., $(\Omega_t, \mathcal{F}_t, \mathcal{P}_t) \xrightarrow{T} (\Omega, \mathcal{F}, \mathcal{P})$. But these are not enough yet.

Definition 2.1: $\{Y_t\}$ is said to be a (strictly) stationary sequence if, for every k and n, $(Y_0, \cdots, Y_n)$ and $(Y_k, \cdots, Y_{k+n})$ have the same distribution.

Stationarity actually means that the time series have the same distribution under the time transformation T.

Recall that any r.v. X on $(\Omega, \mathcal{F}, \mathcal{P})$ transforms its probability space into a real space, i.e.,

$$(\Omega, \mathcal{F}, \mathcal{P}) \xrightarrow{X} (R, \mathcal{B}, P).$$

By Kolmogorov's extension theorem,

$\tilde{X} = (Y_0, Y_1, Y_2, \cdots)$ transforms the space $(\Omega, \mathcal{F}, \mathcal{P})$ into the sample space $(R^N, \mathcal{B}^N, P)$, i.e.,

$$(\Omega, \mathcal{F}, \mathcal{P}) \xrightarrow{\tilde{X}} (R^N, \mathcal{B}^N, P),$$
$$\tilde{X} : w \in \Omega \longrightarrow (w_0, w_1, \cdots) \in R^N.$$

For simplicity, we assume that

$$(\Omega, \mathcal{F}, \mathcal{P}) = (R^N, \mathcal{B}^N, P).$$

Now, $\omega = (w_0, w_1, \cdots) \in \Omega$.

Then,

$$Y_0(w) = w_0 \text{ and } Y_n(w) = w_n \text{ for } n \geq 1.$$

We now define the shift operator T to be the mapping from $\Omega = R^N$ to $\Omega = R^N$:

$$T : w = (w_0, w_1, \cdots) \to (w_1, w_2, \cdots).$$

Thus,

$$Y_n(w) = w_n = Y_0 \left[(w_n, w_{n+1}, \cdots)\right] = Y_0 \left(T^n w\right).$$

The shifts T^k define operators on r.v. Y on $(\Omega, \mathcal{F}, \mathcal{P})$ by

$$(T^k Y)(w) = Y \left[T^k(w)\right].$$

Thus, it follows that

$$Y_n = TY_{n-1} = \cdots = T^{n-1} Y_0$$

and

$$(Y_0, \cdots, Y_n) \xrightarrow{T} (Y_1, \cdots, Y_{n+1}) \xrightarrow{T} \cdots \xrightarrow{T} (Y_k, \cdots, Y_{k+n}),$$

i.e., the time transformation is the shift operator in mathematics.

For $\forall A \in \mathcal{F}$, let $B = T^{-1}A$. Then,

$$\begin{aligned}
P(A) &= P(w \in A) \\
&= P[(w_0, w_1, \cdots) \in A] \\
&= P[w : (Y_0, Y_1, \cdots)(w) \in A] \\
&= P[w : (Y_1, Y_2, \cdots)(w) \in A] \ \ \text{(stationarity)} \\
&= P[w : (w_1, w_2, \cdots) \in A] \\
&= P[w : T(w_0, w_1, \cdots) \in A] \\
&= P(B).
\end{aligned}$$

So we call the shift operator a measure preserving map.

Thus, the stationarity time series $\{Y_t\}$ are determined by the shift operator and Y_0.

Let's go back to $\{r_t\}$. Assume that $\{r_t\}$ is stationary. Then,

$$\mu = \int_{\Omega} r_t d\mathcal{P}$$

is the space average or population average.

When r_t is i.i.d., each observation of r_t can be treated as a sample point from the space. Thus,

$$\mu \approx \frac{1}{n} \sum_{t=1}^{n} r_t. \tag{1}$$

When r_t is not i.i.d., r_t is still a sample point from the space. Thus, we still can count the time average $1/n \sum_{t=1}^{n} r_t$ as some kind of the space average. The difference is that the sample r_t's effect each other, not independent.

How to guarantee that (1) still holds?

In the shoe example, we cam draw a non i.i.d. sample in the following way:

you randomly and continuously walk in China and get one sample point at 6:00 pm each day.

The sample foot size x_t depends on the region where you are on the t-th day, while the region on the t-th day depends on your place on the $(t-1)$-th day. Thus, x_t and x_{t-1} are not independent. After n days, we can get a sample mean. I claim that

$$\mu \approx \frac{1}{n} \sum_{t=1}^{n} x_t \ .$$

Do you believe this?

A condition for this is that you can walk to everywhere or you cannot limit yourself to a region.

i.e., the time transformation T cannot always transform yourself from the region A to region A, for $\forall A$ in China.

Definition 2.2. The shift operator T on $(\Omega, \mathcal{F}, \mathcal{P})$ is said to be ergodic if there is not any $A \in \mathcal{F}$ such that

$$T^{-1}A = A \,[\text{or } P(T^{-1}A = A) = 1]$$

except for $A = \Omega$ or $\emptyset$.

Theorem 2.1. *(1) If $\{Y_t\}$ be a sequence of i.i.d. r.v.s, then $\{Y_t\}$ is stationary and ergodic.*
(2) Let $g : R^N \to R$ be measurable. If $\{X_t\}$ is stationary and ergodic, then $Y_t = g(X_t, X_{t-1}, \cdots)$, $t = -\infty, \cdots, -1, 0, 1, \cdots$, are stationary and ergodic.

- Markov Chain

  A time series $\{Y_t\}$ on $(\Omega, \mathcal{F}, \mathcal{P})$ is said to be a Markov Chain with respect to the σ-filed $\mathcal{F}_t$ if $Y_n \in \mathcal{F}_n$ and for all $B \in \mathcal{F}$,

$$P\left(Y_{n+1} \in B | \mathcal{F}_n\right) = P\left(Y_{n+1} \in B | Y_n\right) \equiv P(Y_n, B)$$

  where $P(x, B)$ is called the transition probability. If there is a distribution π such that

$$\pi(B) = \int P(x, B) d\pi(x), \text{ for } \forall B \in \mathcal{F},$$

  then π is called the invariant distribution (measure).

  If a Markov Chain $\{Y_t\}$ has an invariant distribution π and $Y_0 \sim \pi$, then $\{Y_t\}$ is stationary and ergodic, see Durrett (1995, page335).

Definition 2.3

A Markov Chain $\{Y_t\}$ is called φ-irreducible if there exists a measure φ on Ω such that, for $\forall A \in \mathcal{F}^+$, we have

$$\sum_{n=1}^{\infty} P^n(x, A) > 0$$

for all $x \in \Omega$, where $\mathcal{F}^+ = \{A : \varphi(A) > 0, A \in \mathcal{F}\}$ and $P^n(x, A)$ is the n-step transition of $\{Y_t\}$.

A set A is called recurrent if

$$E\left[\sum_{n=1}^{m} I(Y_n \in A)\middle| Y_0 = x\right] \to \infty$$

as $m \to \infty$, for all $x \in A$.

An φ-irreducible Markov Chain $\{Y_t\}$ is called recurrent if

$$E\left[\sum_{n=1}^{m} I(Y_n \in A)\middle|Y_0 = x\right] \to \infty$$

as $m \to \infty$, for $\forall x \in \Omega$ and $A \in \mathcal{F}^+$. Otherwise, it is called transient.

If $\{Y_t\}$ is φ-irreducible and recurrent, then there exists a unique (up to constant multipler) invariant measure π, see Page 230 of Meyn and Tweedie (1993).

To make $\pi(\Omega) < \infty$ under which we can get a unique invariant probability measure after a normalization, we need another concept.

A set A is called Harris recurrent if

$$P\left(\lim_{m\to\infty}\sum_{n=1}^{m} I(Y_n \in A) = \infty \middle| Y_0 = x\right) = 1,$$

for $\forall x \in A$. A φ-irreducible chain $\{Y_t\}$ is called Harris recurrent if each set in $\mathcal{F}^+$ is Harris recurrent.

A recurrent chain $\{Y_t\}$ differs only by φ-null set from a Harris recurrent chain.

In the finite state space, recurrence is equivalent to Harris recurrence.

If an irreducible chain $\{Y_t\}$ has an invariant probability measure π, then it is recurrent, see Page 231 of Meyn and Tweedie (1993). Thus, the recurrence is necessary for $\{Y_t\}$ to have an invariant probability measure.

For an irreducible chain $\{Y_t\}$, there exists sets $D_1, \cdots, D_d \in \mathcal{F}$ such that

$$(i) \qquad \text{for } x \in D_i, P(x, D_{i+1}) = 1, i = 0, \cdots, d-1,$$

$$(ii) \qquad \text{the set } N = \left[\bigcup_{i=1}^{d} D_i\right]^c \text{ is } \varphi\text{-null and transient,}$$

where d is called the period of $\{Y_t\}$, see Page 188 of Meyn and Tweedie (1993). When $d = 1$, $\{Y_t\}$ is called aperiodic.

An irreducible and periodic example is as follows:

	0	1	2	3
0	0	$\times$	$\times$	0
1	0	0	0	$\times$
2	0	0	0	$\times$
3	$\times$	0	0	0

$$\begin{tikzcd} 0 \arrow[r] \arrow[dr] & 1 \arrow[dl] \\ 3 \arrow[u] & 2 \arrow[l] \end{tikzcd} \qquad \begin{tikzcd} D_1 = \{0\} \arrow[r] & D_2 = \{1, 2\} \arrow[d] \\ & D_3 = \{3\} \arrow[ul] \end{tikzcd}$$

Thus, we have three sub-Markov Chain $Y_t^{(k)} \equiv Y_{3t}|_{Y_0 \in D_k}, k = 1, 2, 3$, and each $Y_t^{(k)}$ is aperiodic and has its own invariant probability distribution on $(D_k, D_k \cap \mathcal{F})$.

Thus, we only focus on the irreducible and aperiodic Markov Chain.

Proposition 2.1. *A necessary and sufficient condition for an irreducible $\{Y_t\}$ to be aperiodic is that there exists an $A \in \mathcal{F}^+$ and $B \subseteq A$ with $B \in \mathcal{F}^+$,*

$$P^n(x, B) > 0 \text{ and } P^{n+1}(x, B) > 0$$

for $\forall x \in B$ and some positive integer n.

In the finite space, an irreducible chain is recurrent as long as one state is recurrent. In the general spaces, a similar idea is there. If we can find out a recurrent $A \in \mathcal{F}^+$, then we can claim that the irreducible chain is recurrent, see Page 174 of Meyn and Tweedie (1993).

What kind of sets are most likely recurrent? It is small set.

Defintion 2.4. *A set $C \in \mathcal{F}$ is called a small set if there exists an $m > 0$, and a non-trivial measure ν_m on $\mathcal{F}$ such that, for all $x \in C$, $B \in \mathcal{F}$,*

$$P^m(x, B) \geq \nu_m(B).$$

The small set means that we start in C and can reach to any set in $\mathcal{F}$ uniformly in a finite step m. It just likes the center of a country.

(a). If $\{Y_t\}$ is a φ-irreducible Feller Chain [i.e., for each bounded continuous function g on R, given x, $E\left[g(Y_t)\middle|Y_{t-1} = x\right]$ is also continuous , see Feigin and Tweedie (1985)], then any relatively compact set A is small if $\varphi(A) > 0$.

(b). If $\{Y_t\}$ is μ_m-irreducible and aperiodic and takes the following form:

$$Y_t = T(Y_{t-1}) + e_t, \quad T : R^m \to R^m,$$

where $Y_t \in R^m$ and $(R^m, \mathcal{B}^m, \mu_m)$ is Lebesgue measure, then any non-null compact set is small if one of the following conditions holds.

(i) $\{e_t\}$ is i.i.d., the marginal distribution is absolutely continuous and ha
a positive pdf over R^m.

(ii) $e_t = (\eta_t, 0, \cdots, 0)'$ with η_t's i.i.d. each having an absolutely continuou
distribution and positive pdf everywhere in R.

Small set is the so-called petite set and a petite set is a small set if $\{Y_t\}$ is φ-irreducible and aperiodic.

To check if a small set C is recurrent, we use a function $V(x)$ to "test" it. In many cases, $V(x)$ can look at as a distance from the starting point x to the "center C". When you hold the Markov Chain to walk around in Ω, the average distance is

$$\int_{\Omega} P(x, dy)V(y).$$

Starting outside of C, it should be less than $V(x)$ when C is recurrent, and larger than $V(x)$ when C is transient.

We now give a little strong drift criterion for the test function $V(x)$. It not only makes sure that $\{Y_t\}$ is Harris recurrent, but also $\pi(\Omega) < \infty$.

Theorem 2.2

Suppose that $\{Y_t\}$ is φ-irreducible and aperiodic. If there exists some small set C, some $b < \infty$, and a non-negative function V finite at some one $x_0 \in \Omega$ such that

$$\text{(i)} \qquad \int P(x,dy)V(y) \leq V(x) - 1 + bI_C(x), \quad x \in \Omega.$$

$$\text{(ii)} \qquad C_V(M) = \{y : V(y) \leq M\} \text{ is a small set for } \forall M > 0,$$

then $\{Y_t\}$ has a unique invariant probability measure π and

$$\sup_{A \in \mathcal{F}} |P^n(x,A) - \pi(A)| \to 0, \tag{2}$$

as $n \to \infty$. If (i) is strengthen as

$$\text{(iii)} \qquad \int P(x,dy)V(y) \leq (1-\beta)V(x) + bI_C(x)$$

for some constants $\beta > 0$ and $b < \infty$, then (ii) hold and

$$\rho^{-n} \sup_{A \in \mathcal{F}} |P^n(x, A) - \pi(A)| \to 0, \tag{3}$$

for some $\rho \in (0, 1)$ as $n \to \infty$.

Under the assumptions of Theorem 2.2, $\{Y_t\}$ is Harris recurrent with an invariant probability, by Theorem 17.1.7 of Meyn and Tweedie (1993), the invariant set $\mathcal{I}$ of $\{Y_t\}$ is trivial to the probability $P(x, \cdot)$ for all $x \in \Omega$. Thus, for $\forall A \in \mathcal{I}$, we have

$$\pi(A) = \int P(x, A)\pi(dx) = 0 \text{ or } 1.$$

Thus, $\{Y_t\}$ is stationary and ergodic.

In the Markov Chain literature, $\{Y_t\}$ is called to be ergodic if (2) holds, which is equivalent to the ergodicity in Theorem 2.1, and is called geometrically ergodic if(3) holds.

If $\{Y_t\}$ is geometrically ergodic, then it is strongly mixing with the rate of convergence ρ^n.

Example 2.1 Suppose $\{\varepsilon_t\}$ are i.i.d. r.v.s with $E|\varepsilon_t|^\gamma < \infty$. Then, the AR(1) model:

$$Y_t = \phi Y_{t-1} + \varepsilon_t \tag{4}$$

has a unique stationary solution if and only if $|\phi| < 1$, and the solution is ergodic with $E|Y_t|^\gamma < \infty$.

Proof.

$$\begin{aligned} Y_t &= \phi Y_{t-1} + \varepsilon_t \\ &= \phi^2 Y_{t-2} + \phi\varepsilon_{t-1} + \varepsilon_t = \cdots \\ &= \phi^m Y_{t-m} + \sum_{i=0}^{m-1} \phi^i \varepsilon_{t-i}. \end{aligned}$$

Let $S_m = \sum_{i=0}^{m-1} \phi^i \varepsilon_{t-i}$. Then, for $\forall m, n > 0$,

$$\begin{aligned} E\left|S_{m+n} - S_m\right|^\gamma &= E\left|\sum_{i=m}^{m+n-1} \phi^i \varepsilon_{t-i}\right|^\gamma \\ &\leq \left(\sum_{i=m}^{m+n-1} |\phi|^{i\gamma}\right) E|\varepsilon_t|^\gamma = O(\rho^n), \end{aligned}$$

for some $\rho \in (0,1)$. By Cauchy criterion, we can show that

$$S_m \to S_\infty, \text{ a.s. and in } L^\gamma, \text{ as } m \to \infty.$$

It is easy to see that $Y_t = S_\infty = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i}$ is a solution of model (4), and it is stationary and ergodic.

To see the uniqueness, we suppose that there is another solution $\{Y_t'\}$ to model (4). Then,

$$\Delta Y_t \equiv Y_t - Y_t' = \phi(Y_{t-1} - Y_{t-1}') = \cdots = \phi^m(Y_{t-m} - Y_{t-m}').$$

Thus,

$$E|\Delta Y_t|^\gamma \leq |\phi|^{\gamma m} \left(E|Y_t|^\gamma + E|Y_t'|^\gamma\right) \to 0,$$

as $m \to \infty$. Hence, $\Delta Y_t = 0$ a.s..

Remark. When $\phi = 1$, we have $Y_t = \sum_{i=1}^t \varepsilon_i + Y_0$. If $E\varepsilon_t^2 = \sigma^2 < \infty$, then

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{[n\tau]} \varepsilon_i \Rightarrow \sigma B(\tau),$$

where $B(\tau)$ is the standard Brownian motion on $C[0,1]$.

When $|\phi| > 1$, we have

$$Y_t = \sum_{i=0}^{t-1} \phi^i \varepsilon_{t-i} + \phi^t Y_0 \to \infty,$$

as $t \to \infty$. It is the explosive case.

If we are allowed to use the future information, we can have a stationary and ergodic solution to model (4) when $|\phi| > 1$.

In fact,

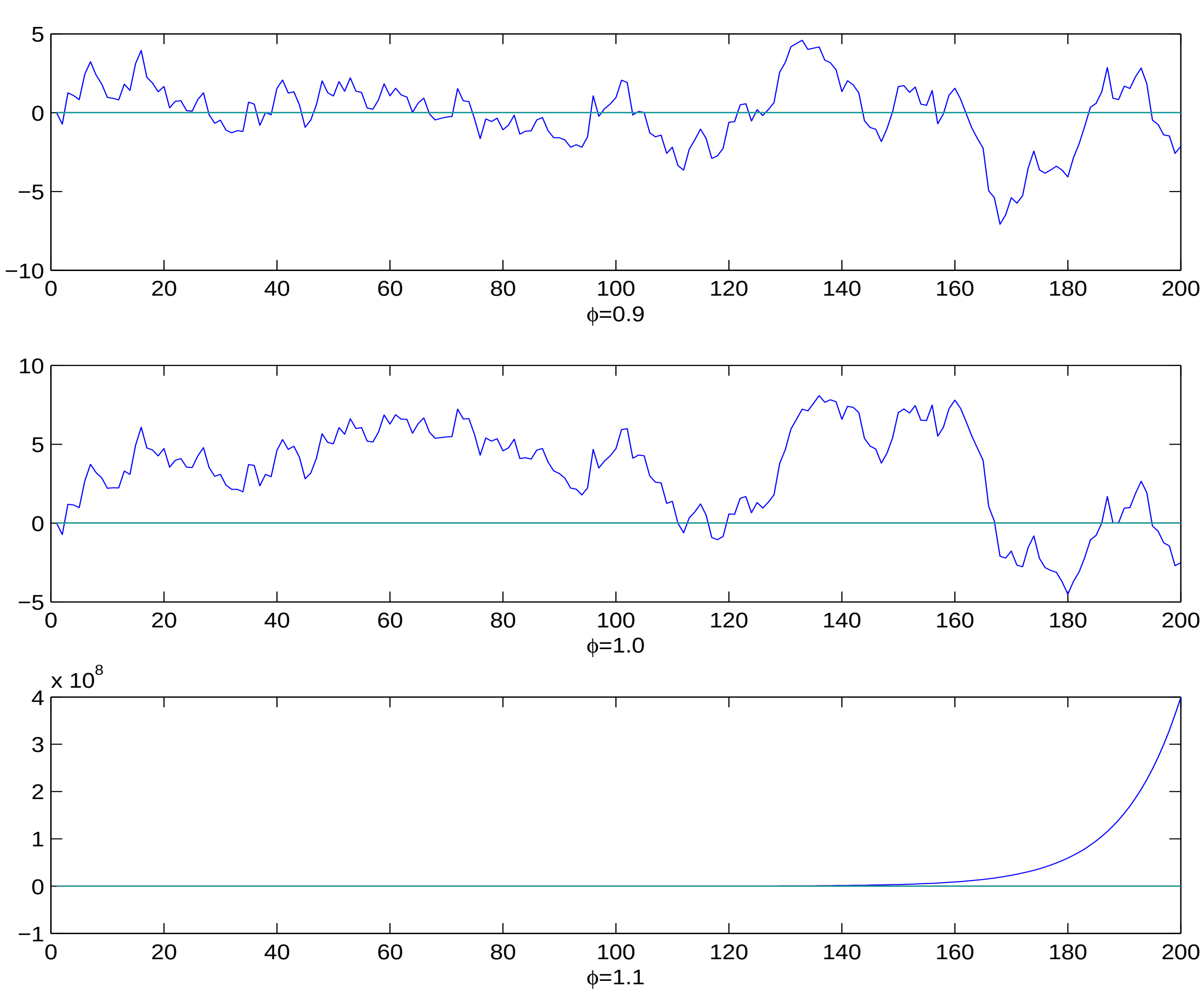
$$Y_t = \phi^{-1} Y_{t+1} - \phi^{-1} \varepsilon_{t+1} = \cdots = \phi^{-m} Y_{t+m} - \left(\sum_{i=1}^{m} \phi^{-i} \varepsilon_{t+i} \right).$$

It follows that

$$S_{mt} = -\sum_{i=1}^{m} \phi^{-i} \varepsilon_{t+i} \rightarrow S_{\infty t} = -\sum_{i=1}^{\infty} \phi^{-i} \varepsilon_{t+i},$$

which is a stationary solution to model (4).

Figure: $\phi = 0.9$, 1.0 and 1.1, $\varepsilon_t \sin N(0, 1)$.



5
0
−5
−10
0 20 40 60 80 100 120 140 160 180 200
φ=0.9
10
5
0
−5
0 20 40 60 80 100 120 140 160 180 200
φ=1.0
x 10^8
4
3
2
1
0
−1
0 20 40 60 80 100 120 140 160 180 200
φ=1.1

Example 2.2. Consider GARCH(1,1) models:

$$\begin{cases} \varepsilon_t = \eta_t \sqrt{h_t}, \\ h_t = \alpha_0 + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}. \end{cases} \tag{5}$$

Model (5) has a unique stationary solution if and only if

$$E \ln |\alpha \eta_t^2 + \beta| < 0, \tag{6}$$

and the solution is ergodic.

Proof. First, if (6) holds, then there exists a $\gamma > 0$ such that

$$\rho \equiv E|\alpha \eta_t^2 + \beta|^\gamma < 1.$$

We now write (5) as

$$
\begin{aligned}
h_t &= \alpha_0 + \left(\alpha\eta_{t-1}^2 + \beta\right) h_{t-1} \\
&= \alpha_0 + \alpha_0 \left(\alpha\eta_{t-1}^2 + \beta\right) + \left(\alpha\eta_{t-1}^2 + \beta\right) \left(\alpha\eta_{t-2}^2 + \beta\right) h_{t-2} \\
&= \cdots \\
&= \alpha_0 \left[1 + \sum_{j=1}^{m} \prod_{i=1}^{j} \left(\alpha\eta_{t-i}^2 + \beta\right)\right] + \prod_{i=1}^{m+1} \left(\alpha\eta_{t-i}^2 + \beta\right) h_{t-m-1}.
\end{aligned}
$$

Let

$$
S_{mt} = \alpha_0 \left[1 + \sum_{j=1}^{m} \prod_{i=1}^{j} \left(\alpha\eta_{t-i}^2 + \beta\right)\right].
$$

For any $m, n > 0$, we have

$$
\begin{aligned}
E\left|S_{m+n,t} - S_{mt}\right|^{\gamma} &= \alpha_0^{\gamma} E \left| \sum_{j=m+1}^{m+n} \prod_{i=1}^{j} \left(\alpha\eta_{t-i}^2 + \beta\right) \right|^{\gamma} \\
&\leq \alpha_0^{\gamma} \sum_{j=m+1}^{m+n} \prod_{i=1}^{j} E\left|\alpha\eta_t^2 + \beta\right|^{\gamma} = O(\rho^n).
\end{aligned}
$$

Thus, by Cauchy criterion, we can show that

$$
S_{mt} \to S_{\infty t} \equiv h_t, \text{ a.s. and in } L^{\gamma}, \text{ as } m \to \infty.
$$

We can verify that h_t is the solution of model (5) and is stationary and ergodic with $Eh_t^{\gamma} < \infty$.

Suppose that we have another solution h_t^* to model (5). Then,

$$\begin{aligned}\Delta h_t &\equiv h_t - h_t^* = (\alpha\eta_{t-1}^2 + \beta)\Delta h_{t-1} = \cdots \\ &= \prod_{i=1}^{m}(\alpha\eta_{t-i}^2 + \beta)\Delta h_{t-m}\end{aligned}$$

and

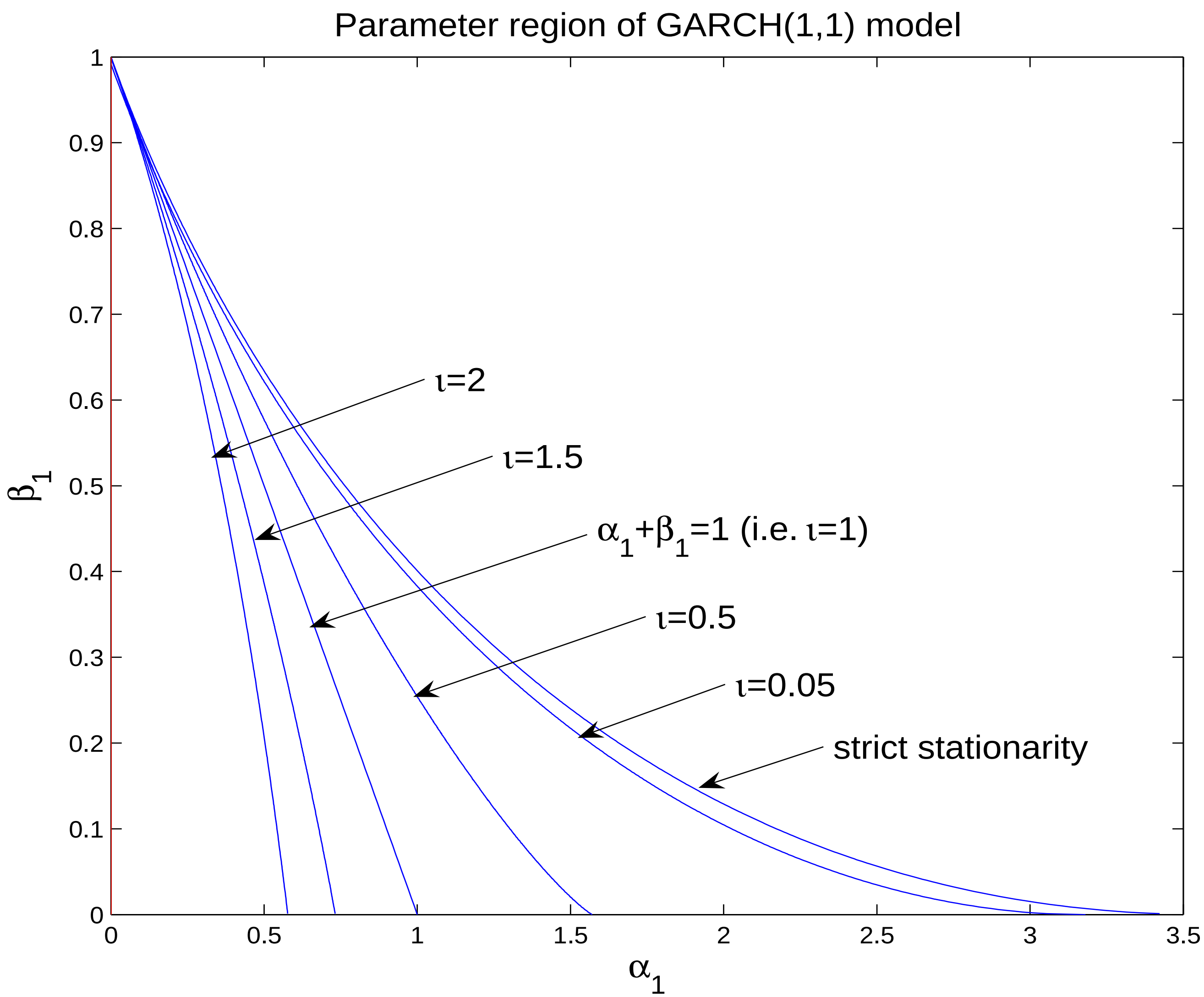
$$E\left|\Delta h_t\right|^\gamma \leq \prod_{i=1}^{m} E\left|\alpha\eta_t^2 + \beta\right|^\gamma E\left|\Delta h_{t-m}\right|^\gamma \leq C\rho^m \to 0, \text{ as } m \to \infty.$$

Thus, $\Delta h_t = 0$ a.s., i.e., $h_t = h_t^*$ a.s..

When $E\ln|\alpha\eta_t^2+\beta|\geq 0$, we have

$$\begin{aligned} h_t &\geq \alpha_0 \max_{1\leq j\leq m} \prod_{i=1}^{j} \left(\alpha\eta_{t-i}^2+\beta\right) \\ &= \alpha_0 e^{\max_{1\leq j\leq m} \sum_{i=1}^{j} \ln\left(\alpha\eta_{t-i}^2+\beta\right)} \longrightarrow \infty, \text{ a.s., as } m \to \infty. \end{aligned}$$

Thus, there is not any solution to model (2.5) when $E\ln|\alpha\eta_t^2+\beta|\geq 0$.



Parameter region of GARCH(1,1) model
ι=2
ι=1.5
α1+β1=1 (i.e. ι=1)
ι=0.5
ι=0.05
strict stationarity
β1
α1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5

Example 2.3 Consider the double AR(1) models:

$$Y_t = \phi Y_{t-1} + \eta_t \sqrt{w + \alpha Y_{t-1}^2}, \tag{7}$$

where $\{\eta_t\}$ are i.i.d. symmetric r.v.s with a continuous and positive density function $f(x)$ on R and $E\eta_t^2 = 1$, $\phi \in R$, w, $\alpha > 0$ and Y_0 is independent of $\{\eta_t\}$. If and only if

$$E \ln |\phi + \sqrt{\alpha}\eta_t| < 0, \tag{8}$$

$\{Y_t\}$ is geometric ergodic and hence has a unique stationary distribution and is strongly mixing with geometric rate of convergence.

Proof. It is easy to see that Y_t is a Markov Chain on $(R, \mathcal{B}, \nu)$ and it is a ν-irreducible Feller Chain since the transition probability function

$$P(x, A) = \int_A \frac{1}{\sqrt{w + \alpha x^2}} f\left(\frac{z - \phi x}{\sqrt{w + \alpha x^2}}\right) dz, \text{ for } A \in \mathcal{B},$$

is strictly positive and continuous, where ν is the Lebesgue measure on $(R, \mathcal{B})$.

This implies that the set $[-M, M]$ is small for $\forall M > 0$. Thus, we only need the drift condition.

Let $h(u) = E\left|\phi + \sqrt{\alpha}\eta_t\right|^u$. Since $h'(0) < 0$, $\exists k > 0$ such that $h(u) < 1$ for every $u \in (0, k)$. Let $\delta \in (0, \min(k, 2))$ and

$$g(x) = 1 + |x|^\delta.$$

Since $f(x)$ is symmetric, we assume that $\phi > 0$ without loss of generality. Note that

$$
\begin{aligned}
E\left[g(Y_t)\middle|Y_{t-1}=x\right] &= 1+E|\phi x+\eta_t\sqrt{w+\alpha x^2}|^\delta \\
&= 1+\left(\frac{w}{\alpha}+x^2\right)^{\frac{\delta}{2}} E\left|\frac{\phi x}{\sqrt{\frac{w}{\alpha}+x^2}}+\sqrt{\alpha}\eta_t\right|^\delta .
\end{aligned}
$$

Since

$$
\frac{\left(\frac{w}{\alpha}+x^2\right)^{\frac{\delta}{2}} E\left|\frac{\phi x}{\sqrt{w/\alpha+x^2}}+\sqrt{\alpha}\eta_t\right|^\delta}{|x|^\delta E\left|\phi+\sqrt{\alpha}\eta_t\right|^\delta} \to 1, \text{ as } x\to\infty,
$$

we have

$$
\begin{aligned}
E\left[g(Y_t)\middle|Y_{t-1}=x\right] &= 1+[1+o(1)]|x|^\delta E\left|\phi+\sqrt{\alpha}\eta_t\right|^\delta \\
&= g(x)+|x|^\delta\left\{[1+o(1)]E\left|\phi+\sqrt{\alpha}\eta_t\right|^\delta-1\right\} \\
&\le g(x)+|x|^\delta(2C-1),
\end{aligned}
$$

as $x \geq M$ and $0 < 2C < 1$. Thus,

$$\begin{aligned} E\left[g(Y_t)\middle|Y_{t-1} = x\right] &\leq g(x) - g(x)\frac{|x|^\delta(1-2C)}{1+|x|^\delta} \\ &\leq g(x) - \frac{g(x)(1-2C)}{2} \\ &\equiv (1-C_0)g(x), \end{aligned}$$

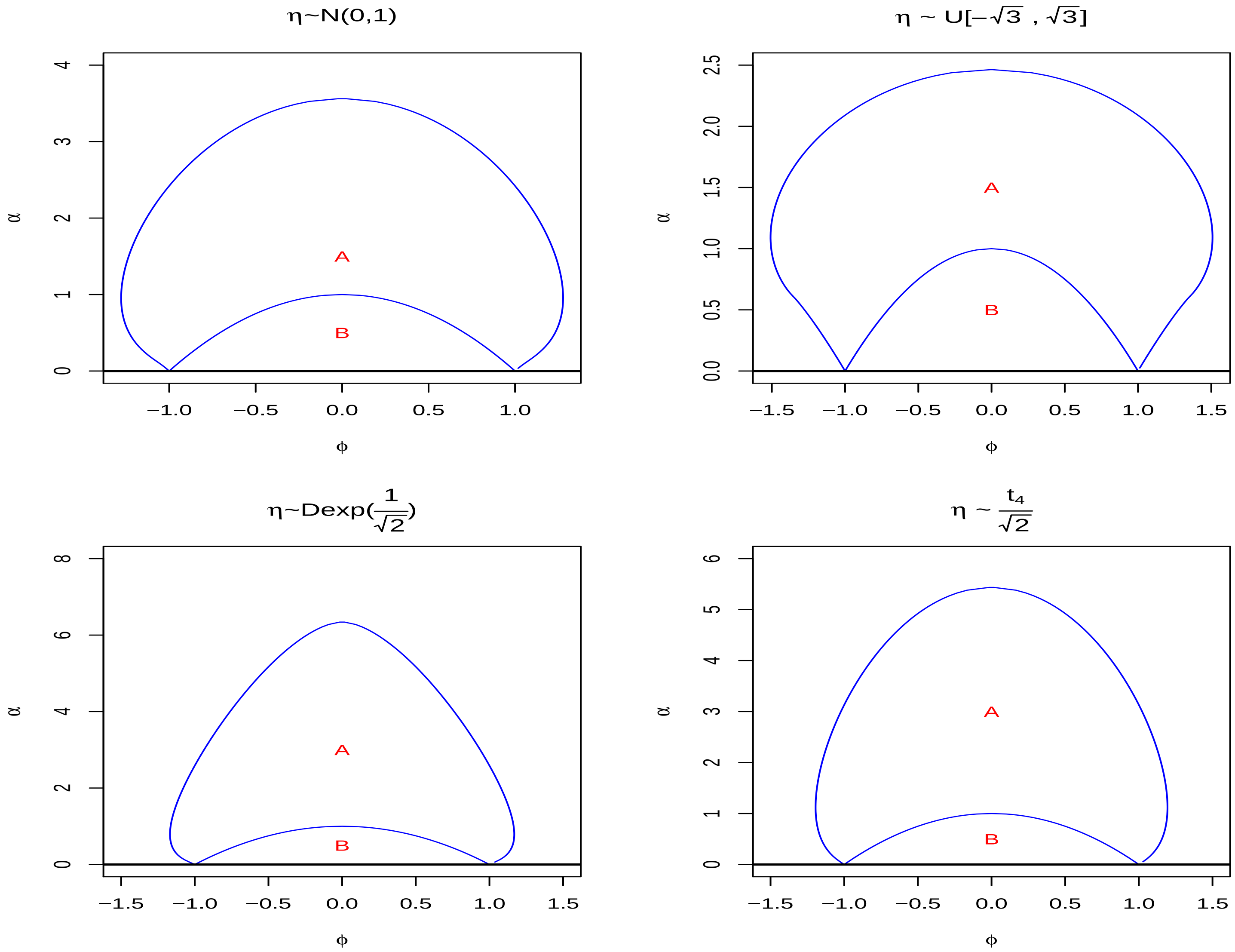
since $|x|^\delta/(1+|x|^\delta) \to 1$. When $|x| \leq M$, it is easy to see that

$$E\left[g(Y_t)\middle|Y_{t-1} = x\right] \leq C_1.$$

Thus, we have

$$E\left[g(Y_t)\middle|Y_{t-1} = x\right] \leq (1-C_0)g(x) + C_1 I_B(x),$$

where $B = [-M, M]$. Now the conclusion follows from Theorem 2.2.


η~N(0,1)
η ~ U[−√3 , √3]
η~Dexp(1/√2)
η ~ t4/√2
A
B
α
ϕ
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
0
1
2
3
4
5
6
8
2.0
2.5

Example 2.4. Consider the TAR(1) models:

$$Y_t = \begin{cases} \phi_1 Y_{t-1} + \varepsilon_t, & \text{if } \; Y_{t-1} > 0, \\ \phi_2 Y_{t-1} + \varepsilon_t, & \text{if } \; Y_{t-1} \le 0, \end{cases} \tag{9}$$

where ε_t is i.i.d. r.v. with mean 0 and each having a strictly positive density $f(\cdot)$ on R. Then, $\{Y_t\}$ is stationary and ergodic if and only if

$$\phi_1 < 1, \phi_2 < 1 \text{ and } \phi_1 \phi_2 < 1. \tag{10}$$

Proof. Note that $\{Y_t\}$ is a Markov Chain with state space $(R, \mathcal{B})$ and its transition density is given by

$$p(x, y) = f\left[y - \phi_1 x I(x > 0) - \phi_2 x I(x \le 0)\right].$$

It is not hard to see that $\{Y_t\}$ is ν-irreducible and aperiodic, where ν is the Lebesgue measure. Since f is strictly positive, any non ν-null compact set $K \in \mathcal{B}$ is small.

From (10), we can select positive constants a and b such that

$$-\frac{a}{b} < \phi_1 < 1 \text{ and } -\frac{b}{a} < \phi_2 < 1.$$

Choosing

$$g(x) = \begin{cases} ax, & \text{if } \ x > 0, \\ -bx, & \text{if } \ x \leq 0. \end{cases}$$

Then,

$$E\left[g(Y_t)|Y_{t-1} = x\right] = \begin{cases} Eg(\phi_1 x + \varepsilon_t), & \text{if } \ x > 0, \\ Eg(\phi_2 x + \varepsilon_t), & \text{if } \ x \leq 0. \end{cases}$$

When $x > 0$ and $\phi_1 < 0$,

$$\begin{aligned} Eg(\phi_1 x + \varepsilon_t) &= ax + aE\left\{\left[(\phi_1 - 1)x + \varepsilon_t\right] I\left(\varepsilon_t > -\phi_1 x\right)\right\} \\ &\quad -bE\left\{\left[(\phi_1 + ab^{-1})x + \varepsilon_t\right] I\left(\varepsilon_t \leq -\phi_1 x\right)\right\} \\ &\leq ax - bE\left\{\left[(\phi_1 + ab^{-1})x + \varepsilon_t\right] I\left(\varepsilon_t \leq -\phi_1 x\right)\right\} + o(x) \\ &< ax - 1, \text{ as } x \to \infty. \end{aligned}$$

When $x > 0$ and $\phi_1 > 0$,

$$\begin{aligned} Eg(\phi_1 x + \varepsilon_t) &\leq ax + aE\left\{\left[(\phi_1 - 1)x + \varepsilon_t\right] I\left(\varepsilon_t > -\phi_1 x\right)\right\} + o(x) \\ &< ax - 1, \text{ as } x \to \infty. \end{aligned}$$

When $x < 0$ and $\phi_2 < 0$,

$$\begin{aligned} Eg(\phi_2 x + \varepsilon_t) &= -bx - bE\left\{\left[(\phi_2 - 1)x + \varepsilon_t\right] I\left(\varepsilon_t \le -\phi_2 x\right)\right\} \\ &\quad + aE\left\{\left[(\phi_2 + a^{-1}b)x + \varepsilon_t\right] I\left(\varepsilon_t > -\phi_2 x\right)\right\} \\ &\le -bx - bE\left\{\left[(\phi_2 - 1)x + \varepsilon_t\right] I\left(\varepsilon_t \le -\phi_2 x\right)\right\} + o(x) \\ &< -bx - 1, \text{ as } x \to \infty. \end{aligned}$$

When $x > 0$ and $\phi_2 > 0$,

$$\begin{aligned} Eg(\phi_2 x + \varepsilon_t) &\le -bx + aE\left\{\left[(\phi_2 + a^{-1}b)x + \varepsilon_t\right] I\left(\varepsilon_t > -\phi_2 x\right)\right\} + o(x) \\ &< -bx - 1, \text{ as } x \to \infty. \end{aligned}$$

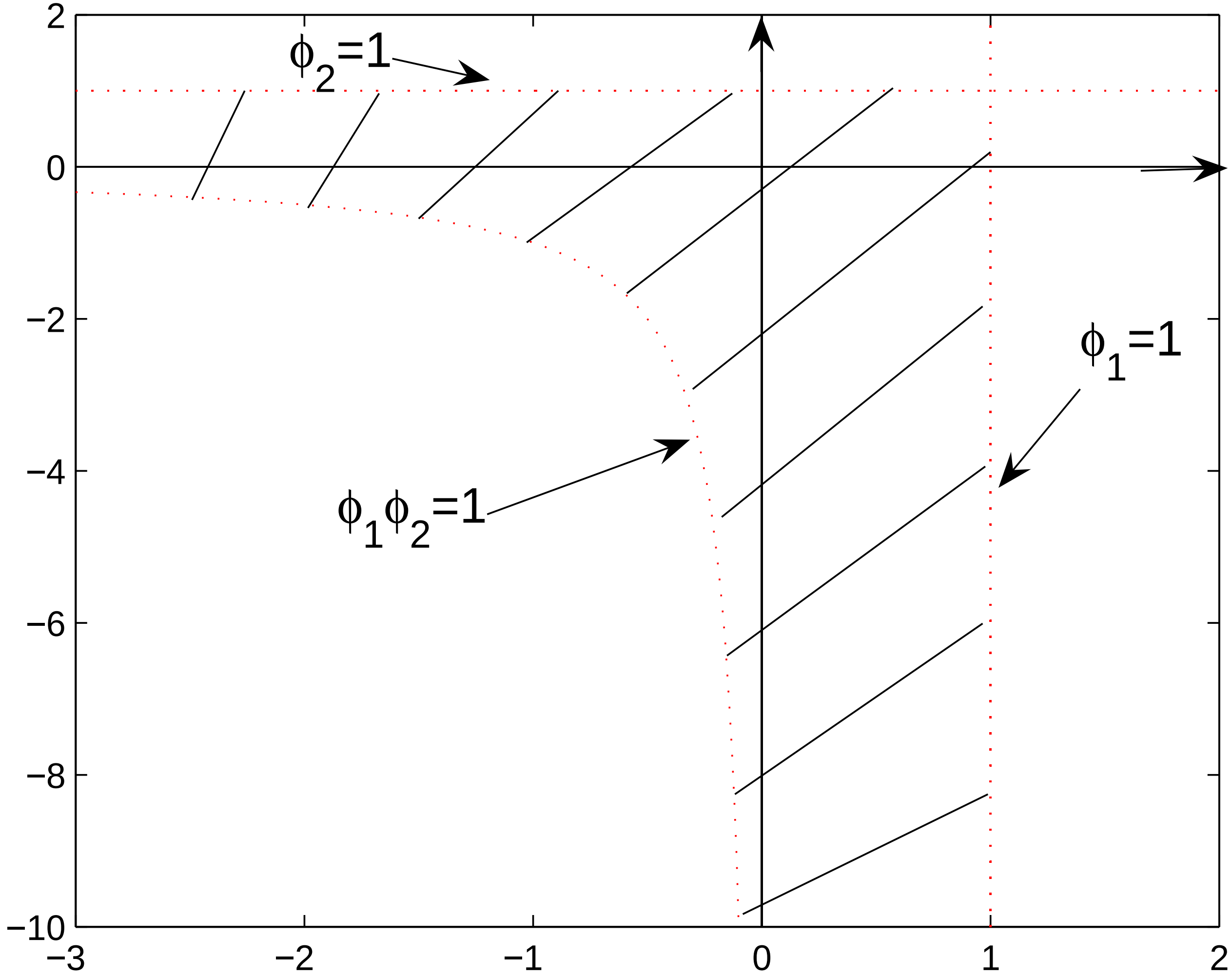
Combining the previous inequalities, we have

$$E\left[g(Y_t|Y_{t-1} = x)\right] \le g(x) - 1, \text{ when } x \notin [-M, M],$$

as M is large enough. It is not hard to see that there exists a constant c such that

$$E\left[g(Y_t|Y_{t-1}=x)\right] \leq g(x) - 1 + cI_{[-M,M]}(x).$$

Thus, there is a unique stationary and ergodic solution to (9).



2
0
−2
−4
−6
−8
−10
−3
−2
−1
0
1
2
φ2=1
φ1=1
φ1φ2=1